Using Taiwan National Health Insurance Database to Model Cancer Insurance

Jack C. Yue
National Chengchi Univ.
Date: Sept. 3, 2014
Email: csyue@nccu.edu.tw

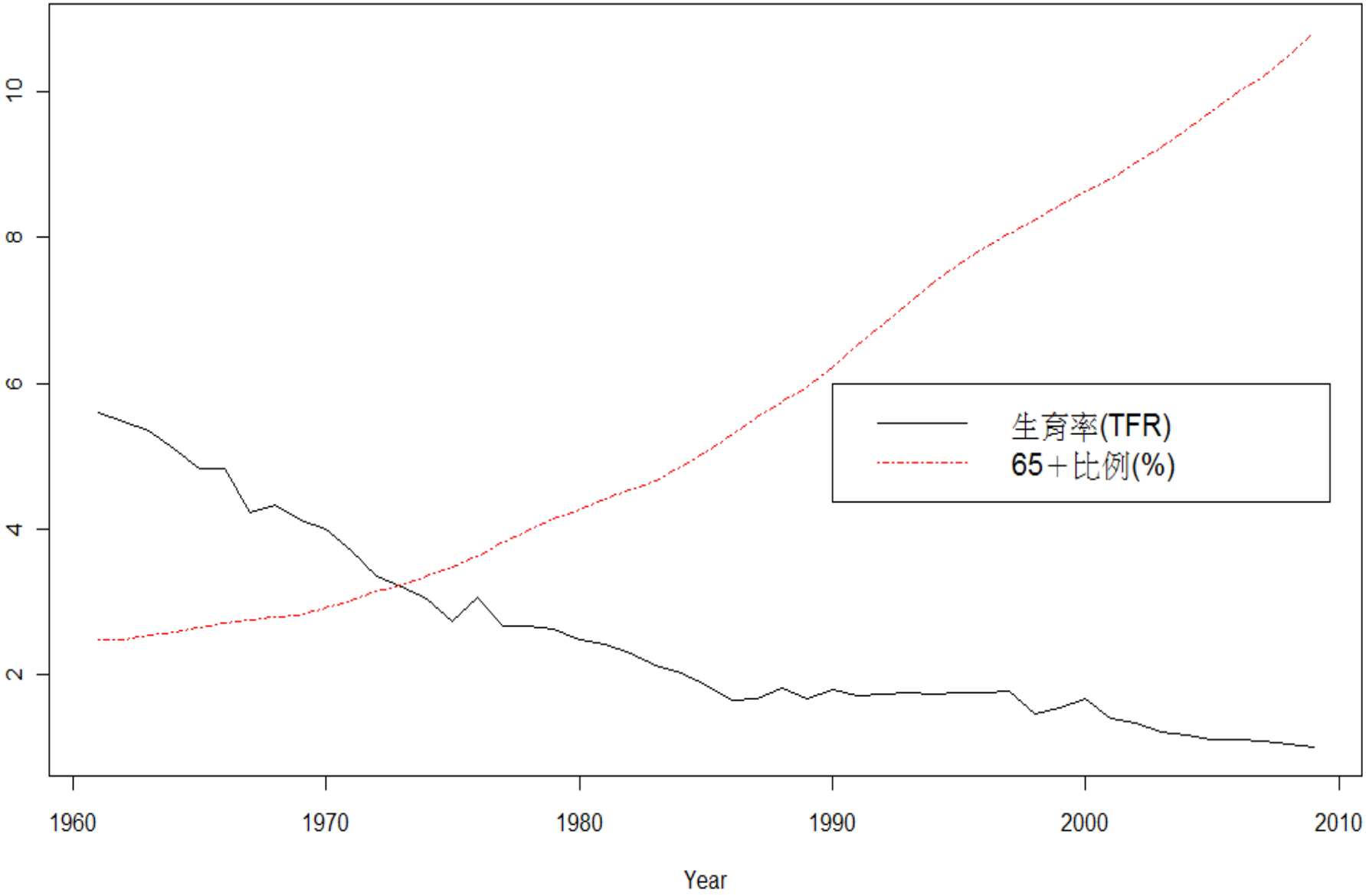
Summary

- Prolonging Life and Cancer
- Taiwan National Health Insurance
- Handling Big Data
- Cancer Incidence and Mortality Rates
- Conclusion and Discussions

Population Trend of 21st Century

- The population aging becomes common to many countries due to prolonging life and lower fertility.
- In addition to **population aging**, **migration** is another common global trend in the 21st century, partly to compensate the shortage of labor force (i.e., not restricted to marriage migration).

Aging and Fertility Trends in Taiwan

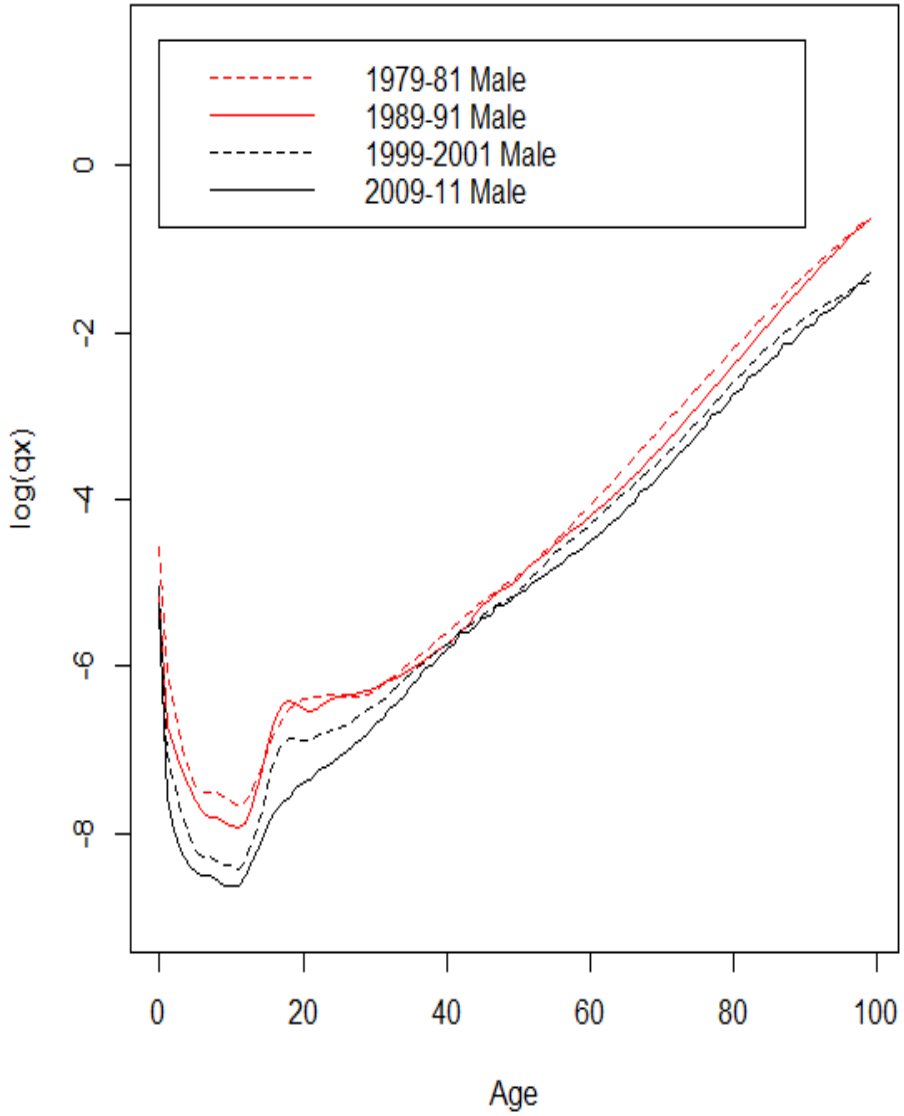


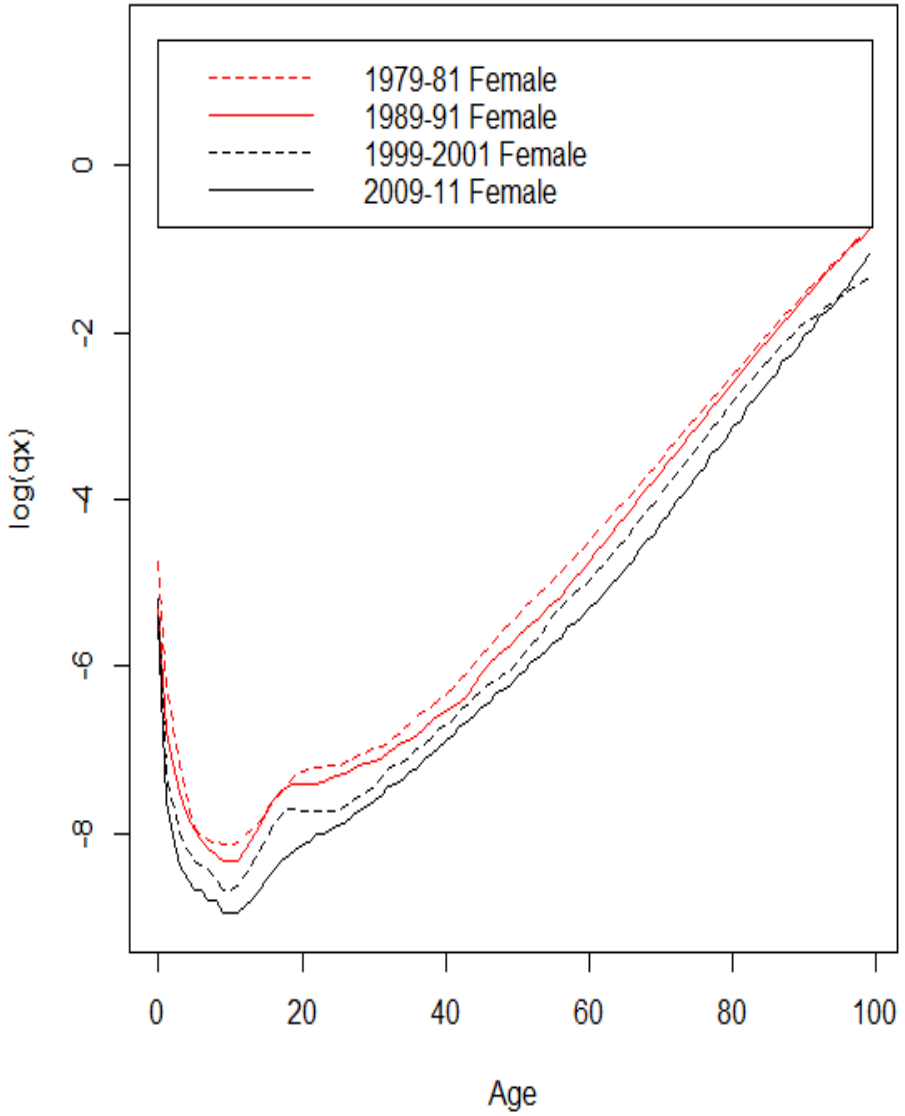
Rapid Population Aging in Taiwan

- Taiwan's 65+ population reached 7% threshold (Ageing Society) in 1993.

→The proportion of 65+ is expected to reach 20% by 2025 (2010 Projection).

→The life expectancy of Taiwan people increases about 0.2~0.3 years annually. (Male and female life expectancies are 75.97 and 82.30 in 2009-2011.)





The Trend of Mortality Rates in Taiwan

Chronic Diseases and Main Death Causes

Taiwan Top 5 Death Causes in 1935

Rank	1	2	3	4	5
Cause	Pneumonia	Enteritis	Parasites Disease	Tuberculosis	Inherence
% Death	21.07%	12.16%	7.12%	6.78%	6.02%

Taiwan Top 5 Causes in 2005

Rank	1	2	3	4	5
Cause	Cancer	Cardiovascular	Heart	Diabete	Accident
% Death	26.79%	9.46%	9.33%	7.56%	6.02%

Sources: Ministry of Interior & Ministry of Health and Welfare

Life Expectancy Eliminated Specified Cause of Death (9 & 10 → 1999-2001 & 2009-11)

Cause of Death	Male		Female	
	10th	9th	10th	9th
Life Expectancy	75.97	73.79	82.32	79.63
Life Expectancy Eliminated Specified Cause				
Cancer	**4.17**	**3.93**	**2.88**	**2.68**
Heart Disease	**1.38**	**1.08**	**1.11**	**0.96**
Accidents	1.08	1.69	0.73	0.80
Cardiovascular Disease	0.88	1.28	0.66	1.28
Pneumonia (Lung Disease)	**0.71**	**0.34**	**0.63**	**0.24**
Diabete	0.66	0.74	0.48	1.36
Chronic Liver Disease	0.64	0.65	0.34	0.35
Lower Respiratory Illnesses	0.53	—	0.33	—
Suicide	0.43	0.32	0.26	0.18
Chronic Kidney Disease	0.34	0.30	0.23	0.42

Higher Medical Expense for the Elderly

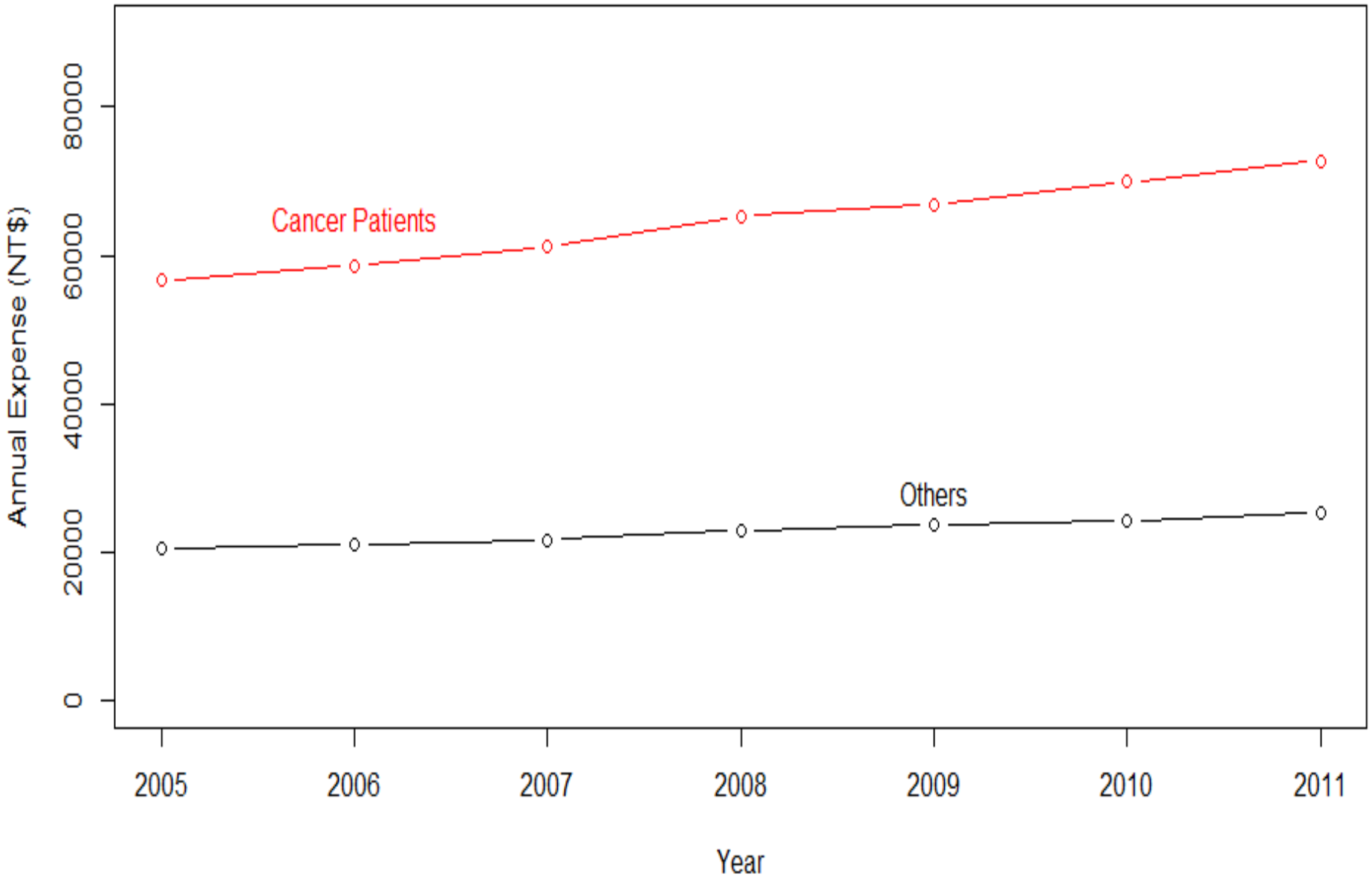
- Elderly spend more in medical expense.

→In 2009, medical expense per person for ages 70~79 is about 6 and 10 times of ages 40~49 and 30~39; The ratio of 65+/0-64 is about 5.

→Ages 65+ spend 29.1% of total expenditure.

Note: The elderly (10%) spend 170 billion NT$, about 1/3 of national health care expenditure.
http://www.nownews.com/2012/03/29/11490-2799490.htm#ixzz2etHeFSoa

Annual Medical Expense of Cancer and other Patients



Study Goal

- The cost of cancer insurance seems to increase and the longevity risk is also a big issue.

→Question: Can we design cancer products?

- Taiwan has maintained complete records of national health care data since 1995.

→Big data?

→Modeling incidence and mortality rates?

Taiwan's National Health Insurance

- Taiwan started the national health insurance (NHI) in 1995, and more than 98% population are covered (excluding oversea works).
- Waiver of copayment is one of the important policy in NHI. In addition to veterans, pregnant women, and people in remote areas, Catastrophic Illness (CI) patients also enjoy the copayment waiver.

Cancer and Catastrophic Illness

- In Taiwan, people with Catastrophic Illness (including cancer) are about 980,000 (4% population) but the medical cost is more than 28% in 2013.

→Cancer is the #1 in Catastrophic Illness, about 48% of people and more than 35% of medical cost in the group of Catastrophic Illness.

Catastrophic Illness Cards

- Evaluating of issuing CI cards is not done by the NHI administration alone.

→e.g., Issuing cancer CI cards requires at least 3 anonymous cancer specialists (committee) from the Chinese Oncology to review medical records and case history.

- Rigorous evaluation has limited the increase of CI cards issued.

Top 10 CI cards at the end of 2012

Rank	Illness	# cards	%	growth
1	癌症**(Cancer)**	463,703	48.2%	7.3%
2	慢性精神病**(Chronically Metal Disease)**	208,325	21.7%	-0.0%
3	需終身治療之全身性自體免疫症候群**(Autoimmune Disease)**	85,473	8.9%	7.3%
4	慢性腎衰竭(尿毒症)必須定期透析治療者**(Kidney)**	71,468	7.4%	4.3%
5	心、肺、胃腸、腎臟、神經、骨骼系統等之先天性畸型及染色體異常**(Congenital Malformations and Chromosomal Abnormalities)**	36,375	3.8%	1.4%
6	先天性新陳代謝疾病**(Congenital Metabolic Diseases)**	14,054	1.5%	5.0%
7	小兒麻痺、腦性麻痺、早產兒所引起之神經、肌肉、骨骼、肺臟等之併發症者(其殘障等級在中度以上者)	12,857	1.3%	1.0%
8	因呼吸衰竭需長期使用呼吸器者**(Human Vegetables)**	11,969	1.2%	3.4%
9	肝硬化症**(Cirrhosis of the Liver Disease)**	10,673	1.1%	-1.8%
10	接受心臟、腎臟及骨髓移植後之追蹤治療**(Organ Transplant)**	10,043	1.0%	7.4%
	Others	36,325	3.8%	1.0%
	Total	961,265	100%	4.6%

Source: Ministry of Health and Welfare

Social and Commercial Insurance

- NHI is jeopardized by population aging and its coverage is likely to further limited.

→Commercial insurance products in Taiwan can serve as the supplements.

- Claiming of commercial products has been deemed to be difficult and criticized.

→If the claim of CI products is based on the insureds are issued with CI cards, it would attract more consumers.

Handling Big Data

- The size and quality of NHI database make data analysis difficult.

→Need to rely on database software and data scientists (e.g., IT experts).

→Data cleaning is a big issue, especially the health care data are from different hospitals.

- Data Discrepancy?

→The death records are not complete in NHI database, and many are even wrong!

	ID		HV		HV_CD	
	# of records	Size (GB)	# of records	Size (GB)	# of records	Size(GB)
1997			877398	0.71	4,821,312	3.34
1998			988,862	0.69	6,115,871	4.24
1999			1,043,563	0.77	7,288,350	5.06
2000			1,109,357	0.91	8,539,223	5.93
2001			1,311,013	1.12	10,062,997	6.98
2002	71,657,446	15.19			11,897,596	7.57
2003	25,130,194	5.33			13,591,212	8.64
2004	25,002,871	5.30			16,287,990	10.36
2005	25,941,824	5.50			17,829,677	11.34
2006	26,336,050	5.58			18,550,542	11.79
2007	26,619,007	5.64	1,518,127	1.43	17,946,211	12.45
2008	26,970,225	5.72	1,583,690	1.46	19,173,919	13.30
2009	27,223,008	5.77	1,785,961	1.65	20,357,193	14.12
2010	27,509,909	7.03	1,923,757	1.83	21,619,442	15.00
2011	27,841,406	5.74	1,996,987	1.90	22,861,178	15.86
合計	66.79 GB		12.47 GB		148.39 GB	

Cleaning the Data

- It is difficult to handle the big data using regular software and the database software (e.g., SQL) is required.
- Data cleaning and exploratory data analysis (EDA) are the key to success. For example,

→For example, more than one databases are available and there exist discrepancy.

Note: ID → Incidence; HV_CD → Mortality

Some examples of database:

Year	ID		HV		HV_CD	
	# persons	# records	# persons	# records	# persons	# records
2007	26,421,498	26,619,007	1,103,431	1,453,483	649,106	17,946,211
2008	26,780,636	26,970,225	1,164,465	1,529,866	678,544	19,173,919
2009	27,067,952	27,223,008	1,276,315	1,733,251	712,828	20,357,173
2010	27,369,795	27,509,909	1,350,786	1,863,254	746,746	21,619,442
2011	27,699,442	27,841,406	1,401,449	1,933,455	779,179	22,861,178

Discrepancy between Different Database?

	HV		HV_CD	
	# of Patients	# of Cases	# of Patients	# of Cases
2007	1,103,431	1,453,483	649,106	17,946,211
2008	1,164,465	1,529,866	678,544	19,173,919
2009	1,276,315	1,733,251	712,828	20,357,173
2010	1,350,786	1,863,254	746,746	21,619,442
2011	1,401,449	1,933,455	779,179	22,861,178

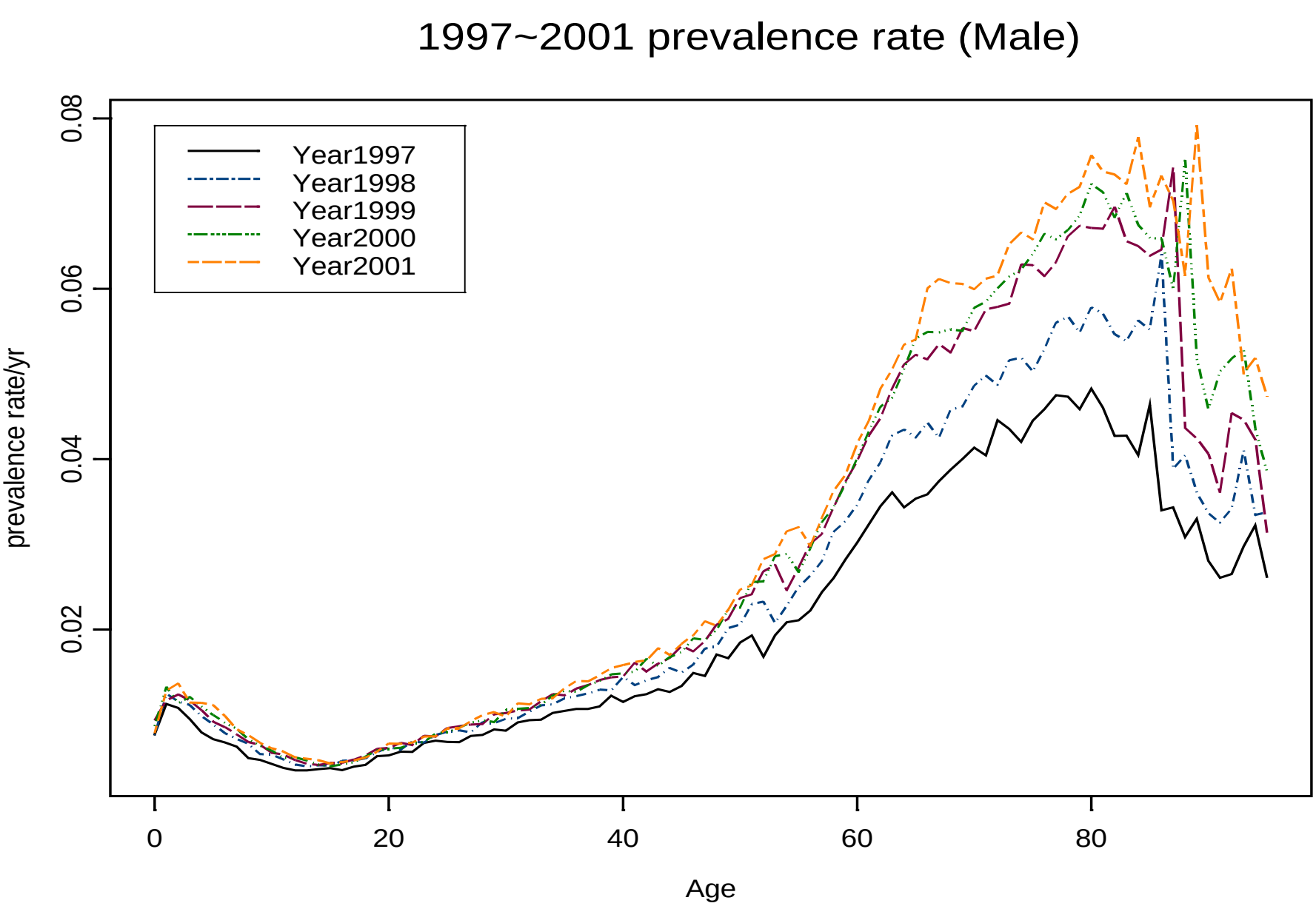
Note: HV and HV_CD contain the list and out-patient visit records of individual CI patient.

Death Note is not good enough…

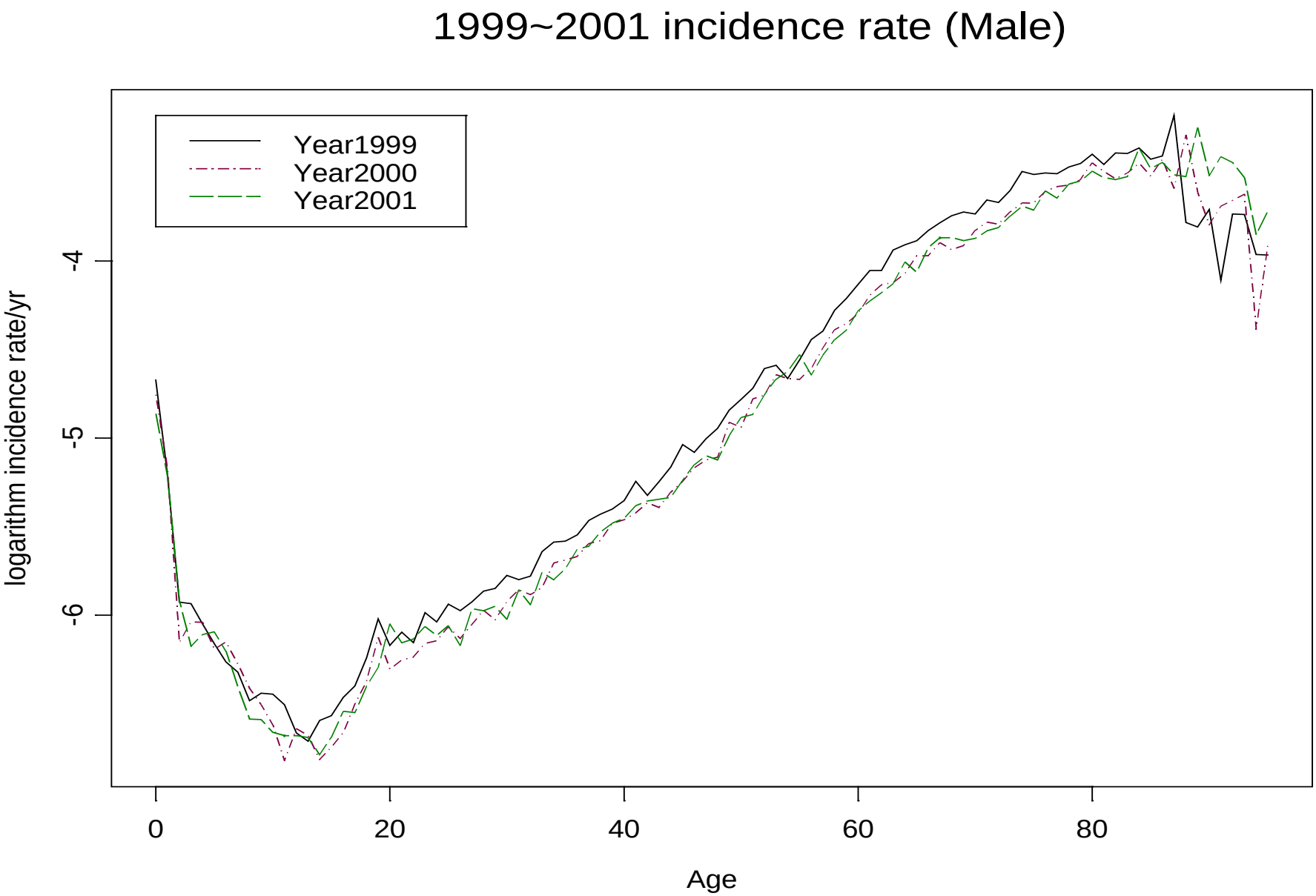
- Death note only covers partial death records and we need to rely on the outpatient visit to judge if patients are still alive.

→e.g., 7,517 persons with “Death Note = Y” in 2007, but 40,306 cancer deaths in official record.

t	Death note at year t	Outpatient Visit at year t	Outpatient Visit at year t+1	Outpatient Visit at year t+2
2007	7,517	1,304	214	0
2008	8,295	1,145	177	0
2009	9,264	898	137	0

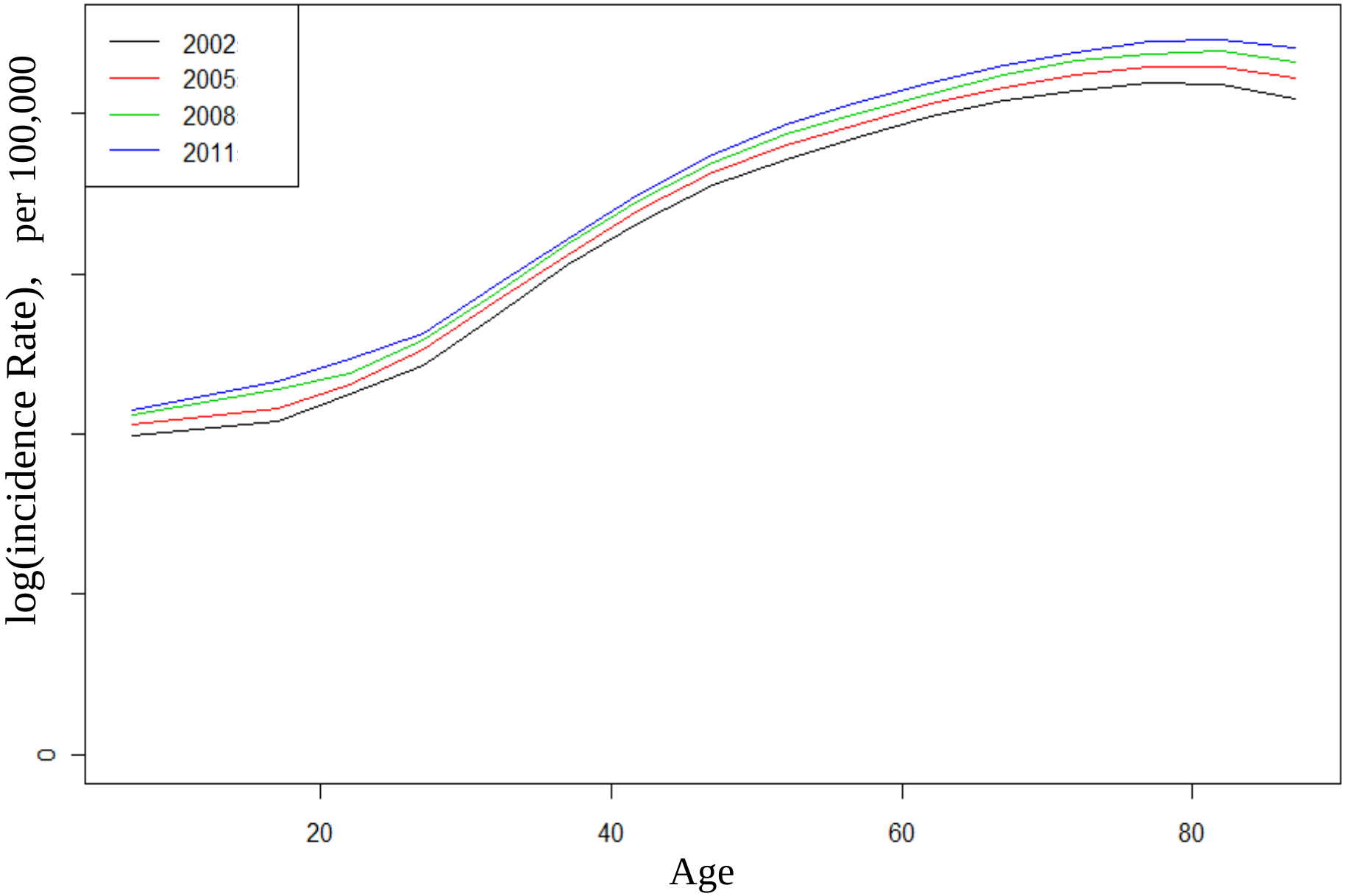


Taiwan's Catastrophic Illness Prevalence Rates



Taiwan's Catastrophic Illness Incidence Rates

Taiwan Age-specific Cancer Incidence Rate



Lee-Carter Model

- The Lee-Carter model (Lee and Carter, 1992) assumes that

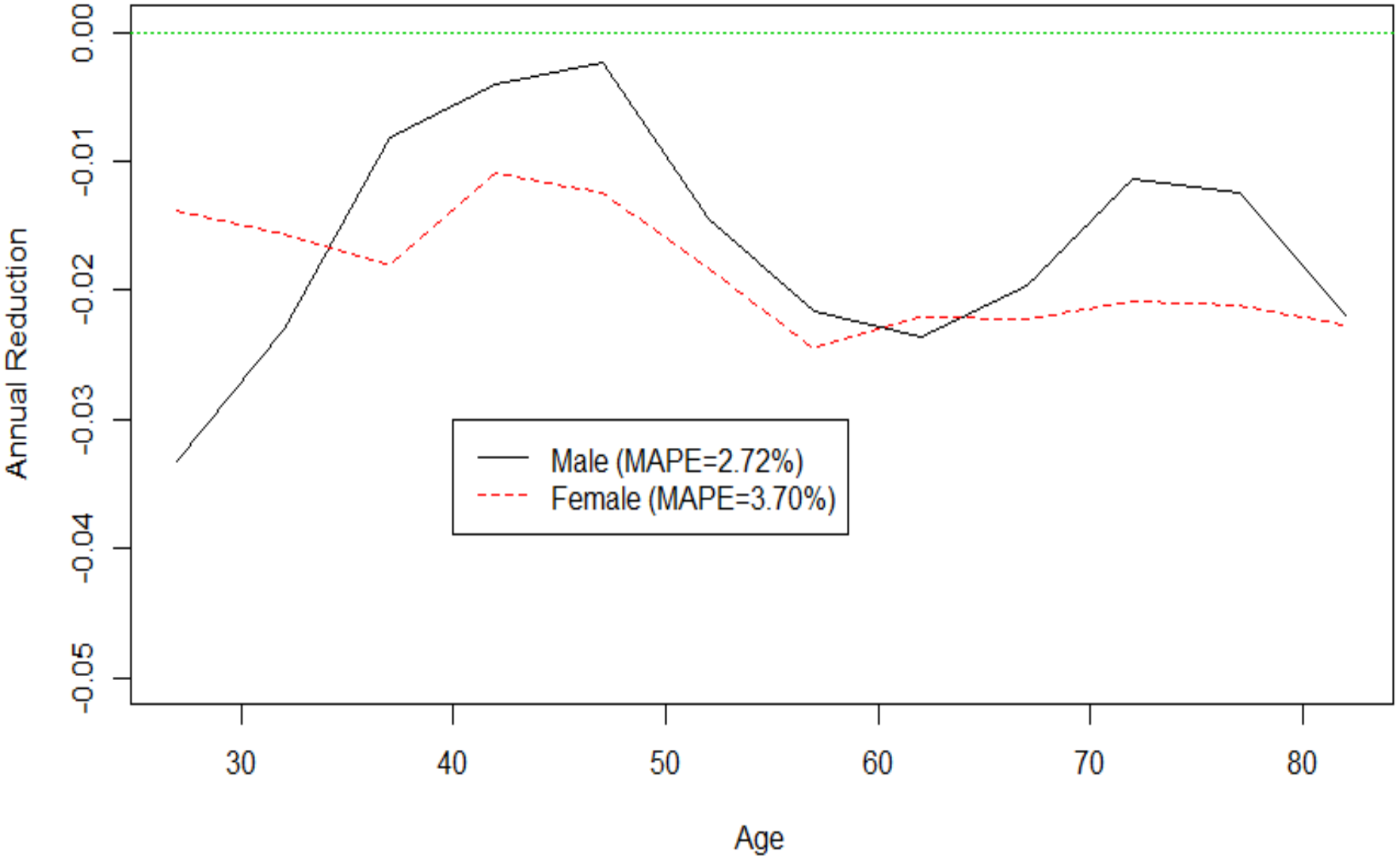
$$\log(m_{x,t}) = \alpha_x + \beta_x \cdot \kappa_t + \varepsilon_{x,t}$$

where x is age, t is time, and α_x, β_x, κ_t are parameters. κ_t is usually a linear function of time.

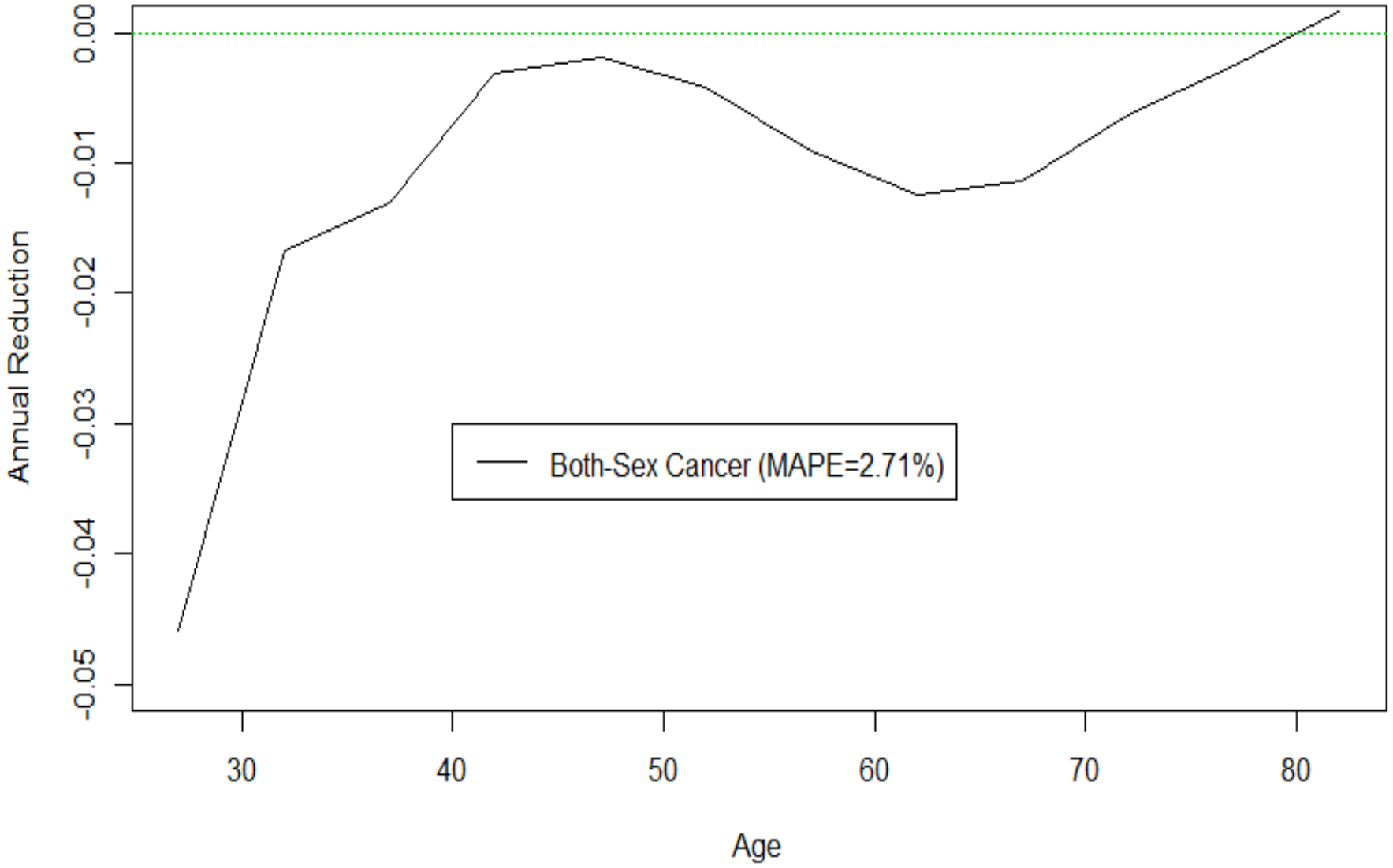
→The LC model often fits well, and it can also be used to fit cancer mortality rates.

Note: Only the data after 1996 are used.

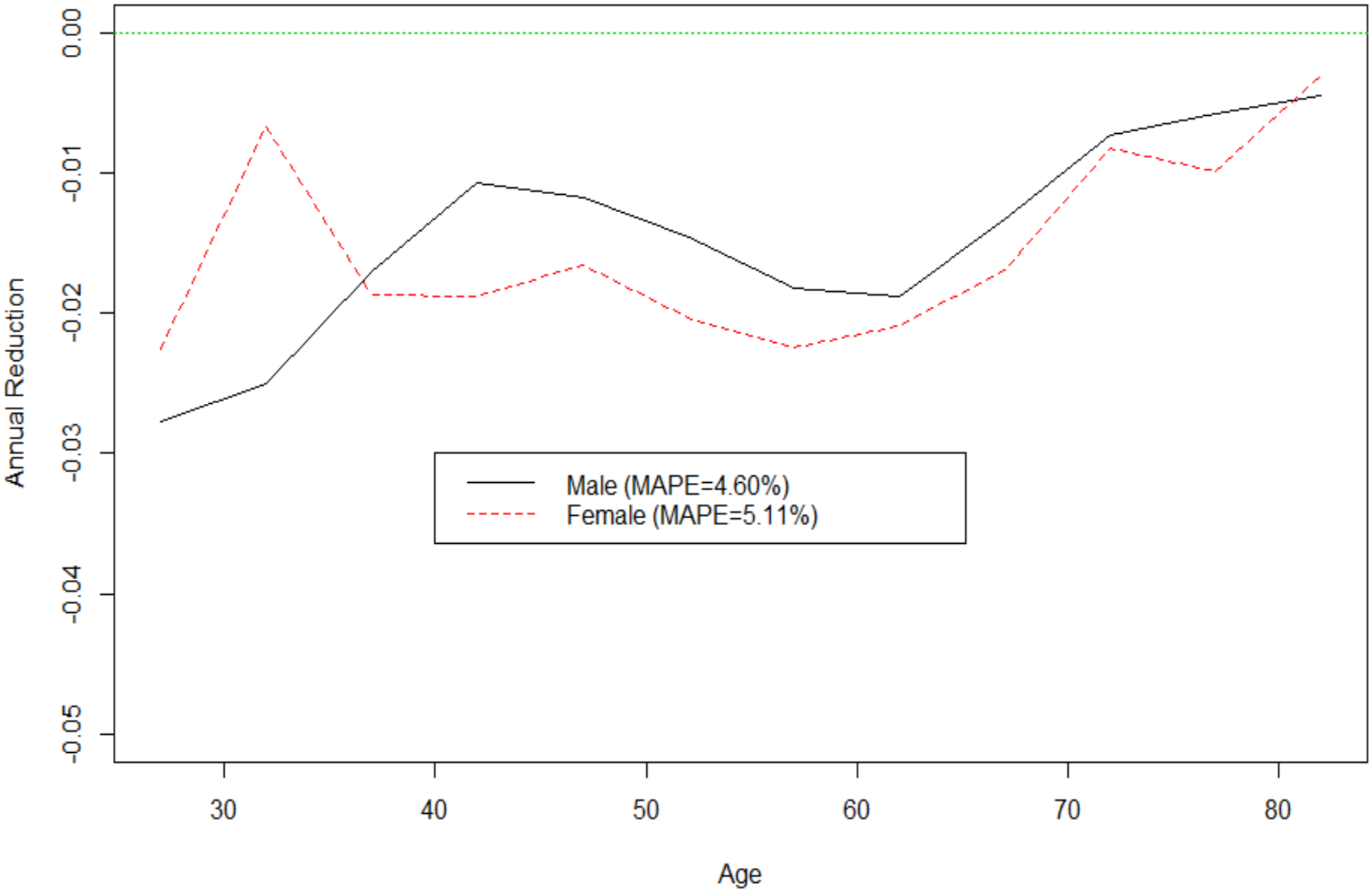
Mortality Reduction under Lee-Carter Model (Total)



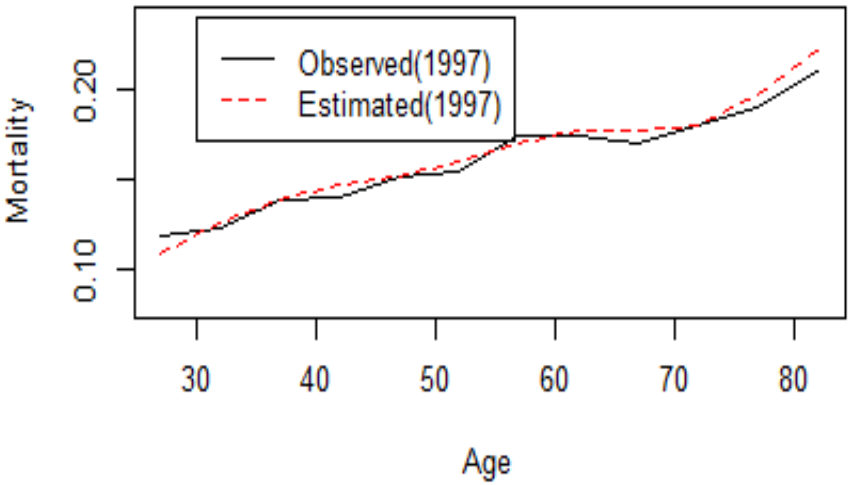
Mortality Reduction under Lee-Carter Model (Cancer)

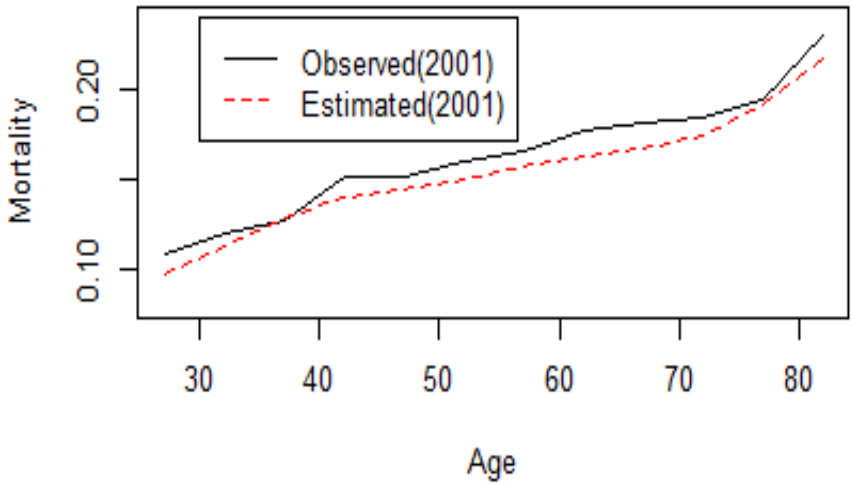


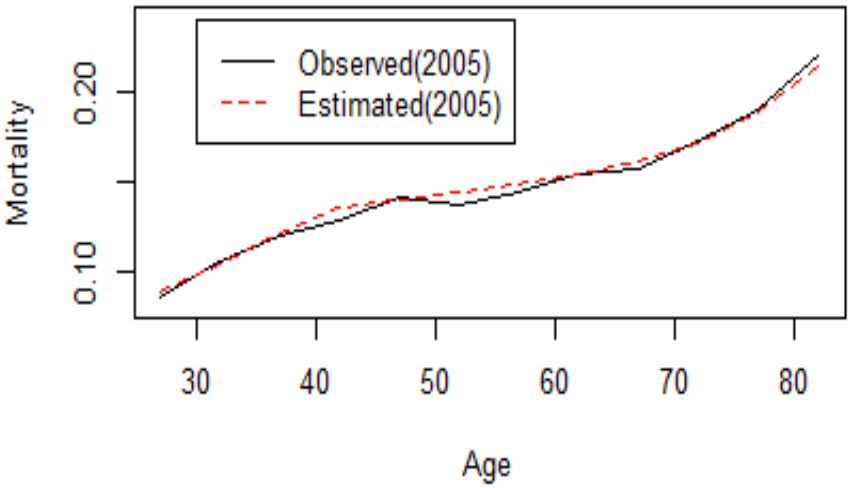
Mortality Reduction under LC Model (Cancer Patients)

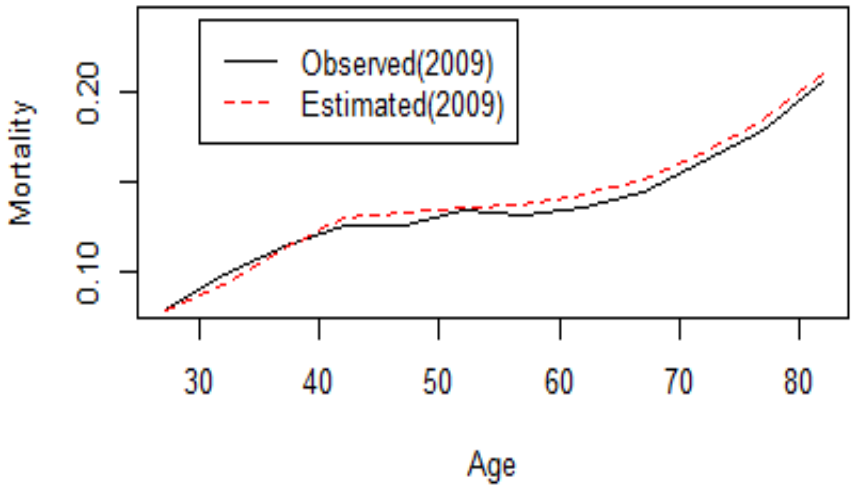


Estimated Result of LC Model (Cancer Male)

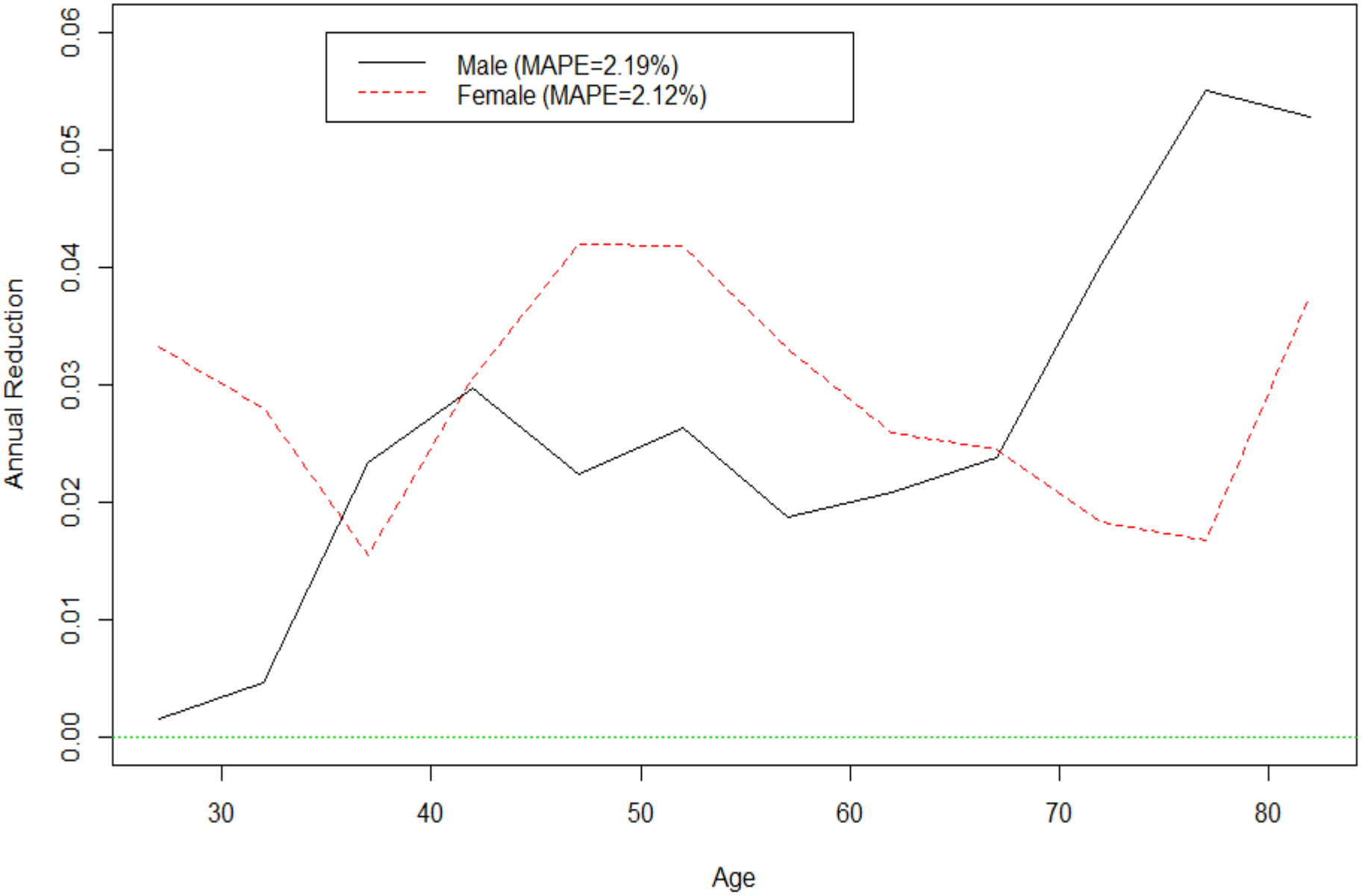




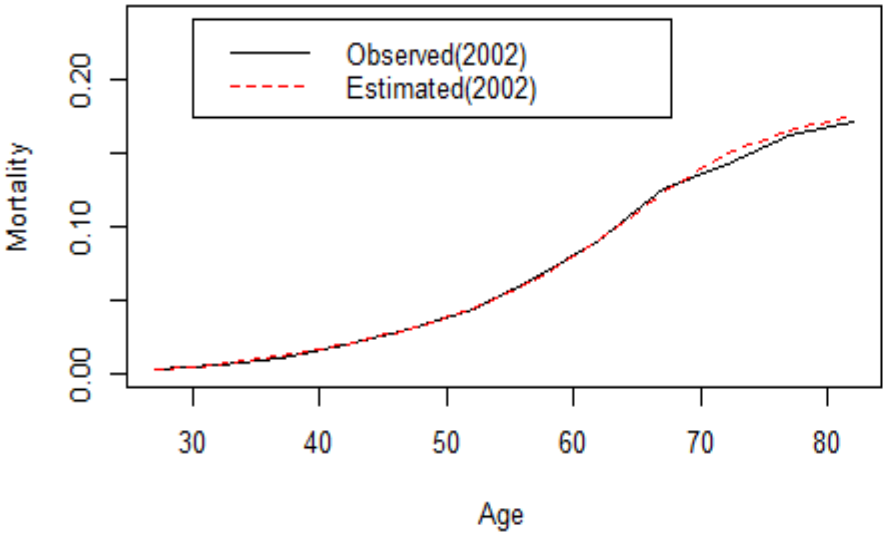


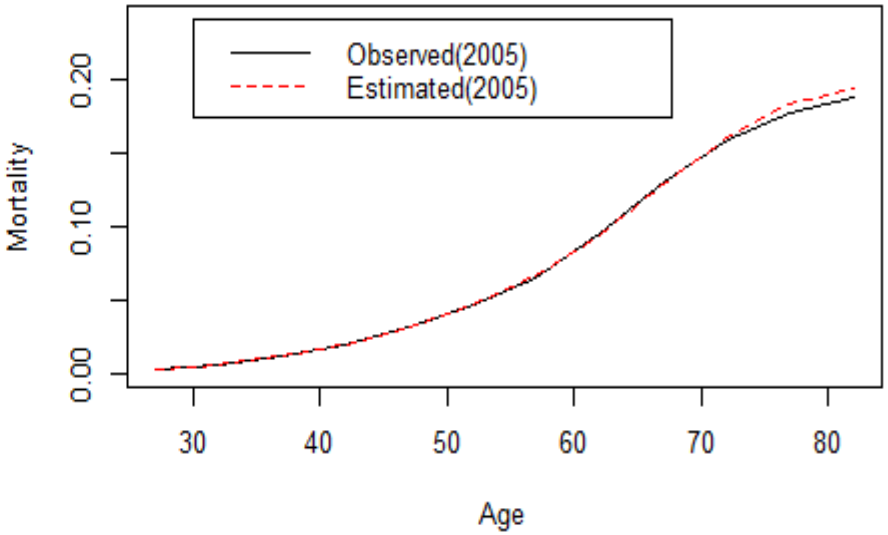


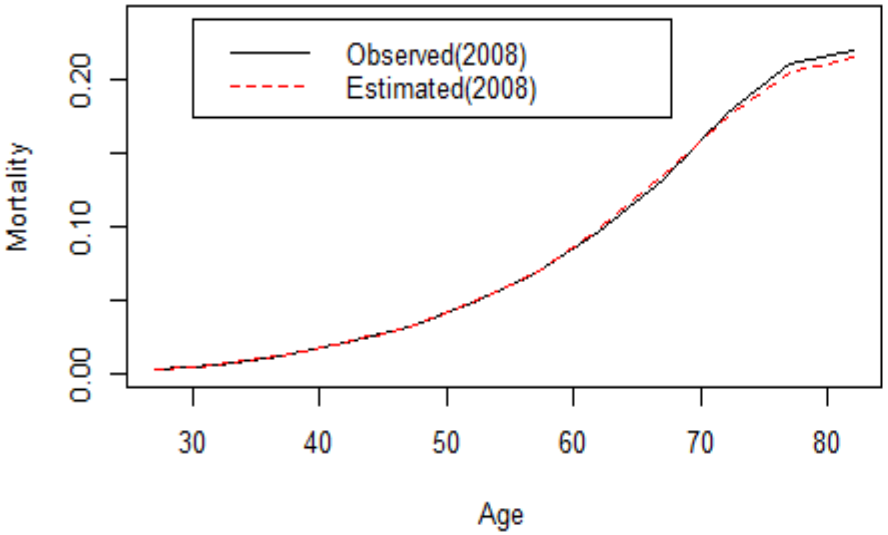
Increasing Incidence under LC Model (Cancer)

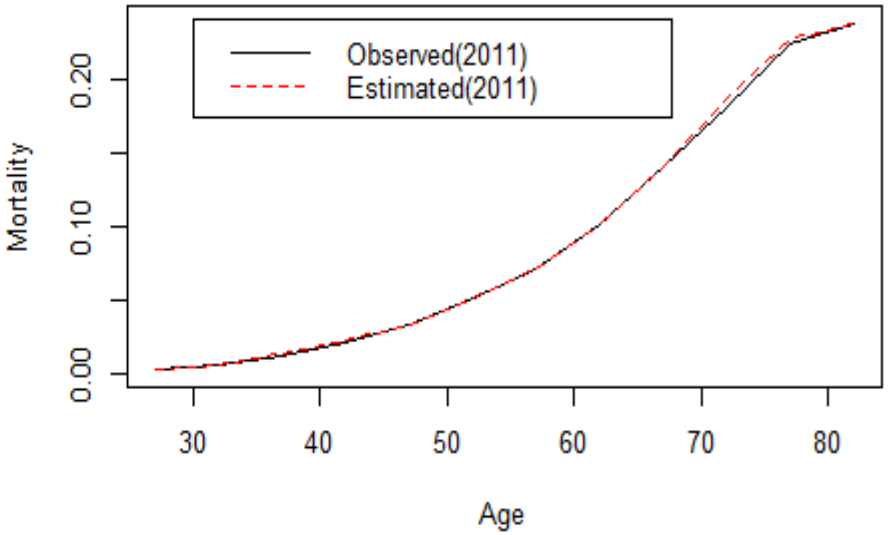


Estimated Result of LC Model (Male Cancer Incidence)









Design Cancer Insurance

- The annual increment in cancer incidence is larger than the decrement in mortality, and cannot be offset by the interest.

→It is especially obvious in the elderly groups.

- Cancer insurance would be expensive if the benefit is payable upon the insured is diagnosed with cancer and the patient is still alive.

→Question: Can we find tools similar to natural hedging to reduce the insurer's risk?

Conclusion

- It takes a lot of efforts to use the database from Taiwan National Health Insurance.

→Knowledge in application fields (e.g., ICD code) and experience accumulation in handling data are important.

- Cancer mortality rates are decreasing, but the cancer incidence rates are increasing.

→Need to consider the longevity risk!

Discussions and Future Study

- Need to accumulate more data and apply survival analysis for survivorship products.

→Medical costs for in- & out-patient visits.

- In addition to actuaries, insurance companies also need experts in big data (& information technology), such as data scientist/statistician.
- Data are an important asset, and regulating the data trading would become necessary.

Thank you
for your Attention!

Q & A